

09/486167

430 Rec'd PCT/PTO 22 FEB 2000

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: UNIVERSITE CATHOLIQUE DE LOUVAIN
Halles Universitaires
- (B) STREET: Place de l' Universite, 1
- (C) CITY: LOUVAIN-LA-NEUVE
- (E) COUNTRY: BELGIUM
- (F) POSTAL CODE (ZIP): B-1348

- (A') NAME: UNIVERSITE DE MONS-HAINAUT
- (B) STREET: Place du Parc 20
- (C) CITY: MONS
- (E) COUNTRY: BELGIUM
- (F) POSTAL CODE (ZIP): B-7000

(ii) TITLE OF INVENTION: PEROXISOME-ASSOCIATED PEPTIDE, NUCLEOTIDE
SEQUENCE ENCODING SAID PEPTIDE AND THEIR USES IN THE
DIAGNOSTIC AND/OR THE TREATMENT OF LUNG INJURIES AND
DISEASES, AND OF OXIDATIVE STRESS-RELATED DISORDERS

(iii) NUMBER OF SEQUENCES: 19

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 805 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 193..681

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GCCAGGAGGC GGAGTGGAAAG TGGCCGTGGG GCGGGTATGG GACTAGCTGG CGTGTGCGCC	60
CTGAGACGCT CAGCGGGCTA TATACTCGTC GGTGGGGCCG GCGGTCAGTC TGCGGCAGCG	120

2' DEOXYRIBONUCLEIC ACID

GCAGCAAGAC GGTGCAGTGA AGGAGAGTGG GCGTCTGGCG GGGTCCGCAG TTTCAGCAGA	180
GCCGCTGCAG CC ATG GCC CCA ATC AAG GTG GGA GAT GCC ATC CCA GCA Met Ala Pro Ile Lys Val Gly Asp Ala Ile Pro Ala	228
1 5 10	
GTG GAG GTG TTT GAA GGG GAG CCA GGG AAC AAG GTG AAC CTG GCA GAG Val Glu Val Phe Glu Gly Glu Pro Gly Asn Lys Val Asn Leu Ala Glu	276
15 20 25	
CTG TTC AAG GGC AAG AAG GGT GTG CTG TTT GGA GTT CCT GGG GCC TTC Leu Phe Lys Gly Lys Gly Val Leu Phe Gly Val Pro Gly Ala Phe	324
30 35 40	
ACC CCT GGA TGT TCC AAG ACA CAC CTG CCA GGG TTT GTG GAG CAG GCT Thr Pro Gly Cys Ser Lys Thr His Leu Pro Gly Phe Val Glu Gln Ala	372
45 50 55 60	
GAG GCT CTG AAG GCC AAG GGA GTC CAG GTG GTG GCC TGT CTG AGT GTT Glu Ala Leu Lys Ala Lys Gly Val Gln Val Val Ala Cys Leu Ser Val	420
65 70 75	
AAT GAT GCC TTT GTG ACT GGC GAG TGG GGC CGA GCC CAC AAG GCG GAA Asn Asp Ala Phe Val Thr Gly Glu Trp Gly Arg Ala His Lys Ala Glu	468
80 85 90	
GGC AAG GTT CGG CTC CTG GCT GAT CCC ACT GGG GCC TTT GGG AAG GAG Gly Lys Val Arg Leu Leu Ala Asp Pro Thr Gly Ala Phe Gly Lys Glu	516
95 100 105	
ACA GAC TTA TTA CTA GAT GAT TCG CTG GTG TCC ATC TTT GGG AAT CGA Thr Asp Leu Leu Leu Asp Asp Ser Leu Val Ser Ile Phe Gly Asn Arg	564
110 115 120	
CGT CTC AAG AGG TTC TCC ATG GTG GTA CAG GAT GGC ATA GTG AAG GCC Arg Leu Lys Arg Phe Ser Met Val Val Gln Asp Gly Ile Val Lys Ala	612
125 130 135 140	
CTG AAT GTG GAA CCA GAT GGC ACA GGC CTC ACC TGC AGC CTG GCA CCC Leu Asn Val Glu Pro Asp Gly Thr Gly Leu Thr Cys Ser Leu Ala Pro	660
145 150 155	
AAT ATC ATC TCA CAG CTC TGA GGCCCTGGGC CAGATTACTT CCTCCACCCCC Asn Ile Ile Ser Gln Leu *	711
160	
TCCCTATCTC ACCTGCCAG CCCTGTGCTG GGGCCCTGCA ATTGGAATGT TGGCCAGATT	771
TCTGCAATAA ACACTTGTGG TTTGCGGAAA AAAA	805

(2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 163 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Ala Pro Ile Lys Val Gly Asp Ala Ile Pro Ala Val Glu Val Phe
 1 5 10 15

Glu Gly Glu Pro Gly Asn Lys Val Asn Leu Ala Glu Leu Phe Lys Gly
 20 25 30

Lys Lys Gly Val Leu Phe Gly Val Pro Gly Ala Phe Thr Pro Gly Cys
 35 40 45

Ser Lys Thr His Leu Pro Gly Phe Val Glu Gln Ala Glu Ala Leu Lys
 50 55 60

Ala Lys Gly Val Gln Val Val Ala Cys Leu Ser Val Asn Asp Ala Phe
 65 70 75 80

Val Thr Gly Glu Trp Gly Arg Ala His Lys Ala Glu Gly Lys Val Arg
 85 90 95

Leu Leu Ala Asp Pro Thr Gly Ala Phe Gly Lys Glu Thr Asp Leu Leu
 100 105 110

Leu Asp Asp Ser Leu Val Ser Ile Phe Gly Asn Arg Arg Leu Lys Arg
 115 120 125

Phe Ser Met Val Val Gln Asp Gly Ile Val Lys Ala Leu Asn Val Glu
 130 135 140

Pro Asp Gly Thr Gly Leu Thr Cys Ser Leu Ala Pro Asn Ile Ile Ser
 145 150 155 160

Gln Leu *

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 780 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Rattus Rattus

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 136..624

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

TGCGTCCTAG GCAGCATAGC CGGATCGGTG CTCCGTGCAT CGGCTACTTG GACGTGCGTG	60
GCAGGCCAGAG CAGGCCGGAA AGGAGCAGGT TGGGAGTGTG GTGGGGCCCG CAGCTTCAGC	120

AGTGCCGCGG	TGACTATGGC	CCCGATCAAG	GTGGGAGACA	CCATTCCCTC	AGTGGAGGTA	180
TTTGRAGGGG	AACCTGGAAA	GAAGGTGAAC	TTGGCAGAGC	TGTTCAAGGA	CAAGAAAGGT	240
GTTTGTTTG	GAGTCCCTGG	GGCATTACCA	CCTGGCTGTT	CCAAGACCCA	TCTGCCTGGG	300
TTTGTGGAGC	AAGCCGGAGC	TCYGAAGGCC	AAGGGAGCAC	AA GTGGTGGC	CTGTCTGAGT	360
GTAAATGATG	YCTTCGTGAC	TGCAGAGTGG	GGTCGAGCCC	ACCAGGCAGA	AGGCAAGGTT	420
CAGCTCCTGG	CTGACCCAC	TGGAGCTTT	GGAAAGGAGA	CAGATTACT	ACTAGATGAT	480
TCTTGGTGT	CTCTCTTGG	GAATCGTCGG	CTAAAAAGGT	TCTCCATGGT	GATAGACAAG	540
GGCGTAGTAA	AGGCACTGAA	CGTGGAGCCG	GATGGCACAG	GCCTCACCTG	CAGCCTGGCC	600
CCCAACATCC	TCTCACAACT	CTGAGGCCCT	GACCAGAATG	TCCTCTGACT	CTCCCATCTC	660
CTCCACCCAG	CTCTGGGCCA	AAGGCCAGT	ACCTCCTTAC	CTGAGGGCCA	CTGGAATGGA	720
ACTTGACAA	TATTTCTGCA	ATAAACAGTT	TAATTTGTGA	AAAAAAAAAA	AAAAAAAAAA	780

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 162 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Rattus Rattus
- (ix) FEATURE:
 - (A) NAME/KEY: Modified-site
 - (B) LOCATION:17
 - (D) OTHER INFORMATION:/product= "Glu or Gly"
- (ix) FEATURE:
 - (A) NAME/KEY: Modified-site
 - (B) LOCATION:63
 - (D) OTHER INFORMATION:/product= "Leu or Pro"
- (ix) FEATURE:
 - (A) NAME/KEY: Modified-site
 - (B) LOCATION:79
 - (D) OTHER INFORMATION:/product= "Ala or Val"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Met	Ala	Pro	Ile	Lys	Val	Gly	Asp	Thr	Ile	Pro	Ser	Val	Glu	Val	Phe
1				5					10				15		

Xaa	Gly	Glu	Pro	Gly	Lys	Lys	Val	Asn	Leu	Ala	Glu	Leu	Phe	Lys	Asp
							20				25			30	
Lys	Lys	Gly	Val	Leu	Phe	Gly	Val	Pro	Gly	Ala	Phe	Thr	Pro	Gly	Cys
							35		40			45			
Ser	Lys	Thr	His	Leu	Pro	Gly	Phe	Val	Glu	Gln	Ala	Gly	Ala	Xaa	Lys
							50		55			60			
Ala	Lys	Gly	Ala	Gln	Val	Val	Ala	Cys	Leu	Ser	Val	Asn	Asp	Xaa	Phe
	65				70					75			80		
Val	Thr	Ala	Glu	Trp	Gly	Arg	Ala	His	Gln	Ala	Glu	Gly	Lys	Val	Gln
							85			90			95		
Leu	Leu	Ala	Asp	Pro	Thr	Gly	Ala	Phe	Gly	Lys	Glu	Thr	Asp	Leu	Leu
							100			105			110		
Leu	Asp	Asp	Ser	Leu	Val	Ser	Leu	Phe	Gly	Asn	Arg	Arg	Leu	Lys	Arg
							115			120			125		
Phe	Ser	Met	Val	Ile	Asp	Lys	Gly	Val	Val	Lys	Ala	Leu	Asn	Val	Glu
							130		135			140			
Pro	Asp	Gly	Thr	Gly	Leu	Thr	Cys	Ser	Leu	Ala	Pro	Asn	Ile	Leu	Ser
	145						150			155			160		
Gln Leu															

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 675 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mouse

(ix) FEATURE:

(A) NAME/KEY: CDS
 (B) LOCATION: 99..588

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TGCTCCGTGC	ATCGACGTGC	TTGGCAGGCA	GAGCAGGCCG	GAAAGAAGCA	GGTTGGGAGT	60
GTGGCGGAGC	CCGCAGCTTC	AGCAGCTCCG	CGGTGACCAT	GGCCCCGATC	AAGGTGGGAG	120
ATGCCATTCC	CTCAGTGGAG	GTATTTGAAG	GGGAACCGGG	AAAGAAGGTG	AACTTGGCAG	180
AGCTGTTCAA	GGGCAAGAAA	GGTGTGTTGT	TTGGAGTCCC	TGGGGCATTT	ACACCTGGCT	240

GTTCTAAGAC CCACCTGCCT GGGTTTGTGG AGCAAGCTGG AGCTCTGAAG GCTAAGGGAG 300
CGCAGGTGGT GGCCTGTCTG AGCGTTAACG ACGTCTTGT GATTGAAGAG TGGGGTCGAG 360
CCCACCAGGC AGAAGGCAAG GTTCGGCTCC TGGCTGACCC CACTGGAGCC TTTGGGAAGG 420
CGACAGACTT ATTATTGGAT GATTCTTGG TGTCTCTCTT TGGGAATCGT CGGCTGAAAA 480
GGTTCTCCAT GGTGATAGAC AACGGCATAG TGAAGGCAG GAACGTGGAG CCAGATGGCA 540
CAGGCCTCAC CTGCAGCCTG GCCCCCCACA TCCTCTCCCA ACTCTGAGGC CCTGGCCAGA 600
TGTCCCTCTGA CTCTCCCAC TCTCCCACCC GGCTCTAGGC CAAAAGGCTC GGTACCTCCT 660
TACTGGGAGC CACGT 675

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 162 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mouse

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met Ala Pro Ile Lys Val Gly Asp Ala Ile Pro Ser Val Glu Val Phe
1 5 10 15

Glu Gly Glu Pro Gly Lys Lys Val Asn Leu Ala Glu Leu Phe Lys Gly
20 25 30

Lys Lys Gly Val Leu Phe Gly Val Pro Gly Ala Phe Thr Pro Gly Cys
35 40 45

Ser Lys Thr His Leu Pro Gly Phe Val Glu Gln Ala Gly Ala Leu Lys
50 55 60

Ala Lys Gly Ala Gln Val Val Ala Cys Leu Ser Val Asn Asp Val Phe
65 70 75 80

Val Ile Glu Glu Trp Gly Arg Ala His Gln Ala Glu Gly Lys Val Arg
85 90 95

Leu Leu Ala Asp Pro Thr Gly Ala Phe Gly Lys Ala Thr Asp Leu Leu
100 105 110

Leu Asp Asp Ser Leu Val Ser Leu Phe Gly Asn Arg Arg Leu Lys Arg
115 120 125

Phe Ser Met Val Ile Asp Asn Gly Ile Val Lys Ala Leu Asn Val Glu
130 135 140

Pro Asp Gly Thr Gly Leu Thr Cys Ser Leu Ala Pro Asn Ile Leu Ser
145 150 155 160

Gln Leu

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 469 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 161..382

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGGTATGGGA CTAGCTGGCG TGTGCGCCCT GAGACGCTCA GCGGGCTATA TACTCGTCGG	60
TGGGGCCGGC GGTCAGTCTG CGGCAGCGGC AGCAAGACGG TGCAGTGAAG GAGAGTGGC	120
GTCTGGCGGG GTCCGCAGTT TCAGCAGAGC CGCTGCAGCC ATGGCCCCAA TCAAGGTTCG	180
GCTCCTGGCT GATCCCCTTG GAAGGAGACA GACTTATTAC TAGATGATTC	240
GCTGGTGTCC ATCTTGGA ATCGACGTCT CAAGAGGTTTC TCCATGGTGG TACAGGATGG	300
CATAGTGAAG GCCCTGAATG TGGAACCAGA TGGCACAGGC CTCACCTGCA GCCTGGCACC	360
CAATATCATC TCACAGCTCT GAGGCCCTGG GCCAGATTAC TTCCTCCACC CCTCCCTATC	420
TCACCTGCC AGCCGTGTGC TGGGGCCCTG CAATTGGAAT GTTGGCCAG	469

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 601 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION:161..514

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGGTATGGGA	CTAGCTGGCG	TGTGCGCCCT	GAGACGCTCA	GCGGGCTATA	TACTCGTCGG	60
TGGGGCCGGC	GGTCAGTCTG	CGGCAGCGGC	AGCAAGACGG	TGCAGTGAAG	GAGAGTGGGC	120
GTCTGGCGGG	GTCCGCAGTT	TCAGCAGAGC	CGCTGCAGCC	ATGGCCCCAA	TCAAGACACA	180
CCTGCCAGGG	TTTGTGGAGC	AGGCTGAGGC	TCTGAAGGCC	AAGGGAGTCC	AGGTGGTGGC	240
CTGCTGAGT	GTTAATGATG	CCTTGATGAC	TGGCGAGTGG	GGCCGAGCCC	ACAAGGCGGA	300
AGGCAAGGTT	CGGCTCCTGG	CTGATCCCAC	TGGGGCCTTT	GGGAAGGAGA	CAGACTTATT	360
ACTAGATGAT	TCGCTGGTGT	CCATCTTGG	GAATCGACGT	CTCAAGAGGT	TCTCCATGGT	420
GGTACAGGAT	GGCATAGTGA	AGGCCCTGAA	TGTGGAACCA	GATGGCACAG	GCCTCACCTG	480
CAGCCTGGCA	CCCAATATCA	TCTCACAGCT	CTGAGGCCCT	GGGCCAGATT	ACTTCCTCCA	540
CCCCCTCCCTA	TCTCACCTGC	CCAGCCCTGT	GCTGGGGCCC	TGCAATTGGA	ATGTTGGCCA	600
G						601

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 604 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION:161..517

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGGTATGGGA	CTAGCTGGCG	TGTGCGCCCT	GAGACGCTCA	GCGGGCTATA	TACTCGTCGG	60
------------	------------	------------	------------	------------	------------	----

TGGGGCCGGC GGTCAGTCTG CGGCAGCGGC AGCAAGACGG TGCAGTGAAG GAGAGTGGC	120
GTCTGGCGGG GTCCGCAGTT TCAGCAGAGC CGCTGCAGCC ATGGCCCCAA TCAAGGTGG	180
AGATGCCATC CCAGCAGTGG AGGTGTTGA AGGGGAGCCA GGGAACAAAGG TGAACCTGGC	240
AGAGCTGTTA AAGGGCAAGA AGGGTGTGCT GTTGAGGTT CCTGGGGCCT TCACCCCTGG	300
ATGTTCCAAG GTTCGGCTCC TGGCTGATCC CACTGGGCC TTTGGGAAGG AGACAGACTT	360
ATTACTAGAT GATTGCTGG TGTCCATCTT TGGGAATCGA CGTCTCAAGA GGTTCTCCAT	420
GGTGGTACAG GATGGCATAG TGAAGGCCCT GAATGTGGAA CCAGATGGCA CAGGCCTCAC	480
CTGCAGCCTG GCACCCAATA TCATCTCACA GCTCTGAGGC CCTGGGCCAG ATTACTTCCT	540
CCACCCCTCC CTATCTCACC TGCCCAGCCC TGTGCTGGGG CCCTGCAATT GGAATGTTGG	600
CCAG	604

(2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2710 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
- (ix) FEATURE:
 - (A) NAME/KEY: exon
 - (B) LOCATION:2516..2710
- (ix) FEATURE:
 - (A) NAME/KEY: exon
 - (B) LOCATION:2074..2135
- (ix) FEATURE:
 - (A) NAME/KEY: exon
 - (B) LOCATION:1932..1970
- (ix) FEATURE:
 - (A) NAME/KEY: exon
 - (B) LOCATION:1728..1859
- (ix) FEATURE:
 - (A) NAME/KEY: exon
 - (B) LOCATION:802..936

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

TCTGTCCCTT AGCGCCCCCG CGGGGGCTTA CCCCATCCCA CTCCATGACC TCCCCCTCCCC 60
CCATGGCGAA TTCCCACCTT TCTGTCTTC ACTCACTTCC TGGAACCGTC CCCAGGGCCT 120
TGGACCTTCC CCCTTCTCCT CCCAACCTT GTGAGACCCC ATTCCCTTTC TACTTCATCC 180
TGCTCTCAAC TTTTGGGCTC CTCAGAGGCC CTCACCCCTG ACTCTCTCTC CCTACCCACT 240
CTGGTCCCCT GAAGCCCTCA AGTACTCTGG GGATGGATCC TTCCCCCTTC AAAAGATTCC 300
TTCTTTGTT CTACACCTCC TGGGTGTAAGG GGCCTGGACA CCCTCCCCCA ACGTTCCACC 360
TGCCGCTGCC CTTCCCTTCC CTCCTCCTGA GGGTGGGACC CTCAGACCTG GCCAAGATCC 420
TCTCCCTCCA TGTTGTCAGG GACTCCTCCT CACCCCCAAA TACAGCCCTC TAGCCCTGT 480
CCATTTATT CCACTCCTT CCTGTAACCT AGACAGCATG TTATGCAACC CTTGCGACA 540
CATGGGAAA CCTTCCCTCC CTTCCTCTGT TGTCACCAAT GGCCCTTAA GAGGAGCAGG 600
GCCACCTTGA AACTTGGAGG ATATGGGTA ACCCAGTGGG AGCGGGCAGG GAGGCCCTT 660
GGAAACTGAC AGGGCTGGAG TATCCTGCTG GGTTTCAGCC CCGGTTCCCTG CAGGCACAGC 720
TGCCAGGCTC TCTGTTCAACC TTCCCTGCCTC TGGTTTGCCTC CGGCTCCCTC ACCCCCCCTTA 780
CCCTGGAGTC CTTCCCTCTA GGTGGGAGAT GCCATCCCAG CAGTGGAGGT GTTGAAGGG 840
GAGCCAGGGA ACAAGGTGAA CCTGGCAGAG CTGTTCAAGG GCAAGAAGGG TGTGCTGTTT 900
GGAGTTCCCTG GGGCCTTCAC CCCTGGATGT TCCAAGGTGA GGCCCTTCCC CTTCTGAAGA 960
TCAGGACCTG GGGATCTTT GTGTTGCTCT TAAGTCCTCC ACATAGTCCT GATAGGACTC 1020
CTAAAAAGCA TTTCAGTGCC ATCACAAAAC AAGTAGAGCT GGGTAGAGCT GGGCGCGGTG 1080
GCTCACGCCT GTAATCCCAG CACTTGGGA GGCAAGGCG GGTGGATCAC GAGGTCAGGA 1140
GTCCAAAACC AGCCTGGCCA AGATGGTGAA ACCCTGTCTC TACTAAAAAT GCAAAAAAAT 1200
CAGCCGGATA TGGTGGCGGG CGCCTGTAAT CCCAGGTATT GGGGAGGCTG AGGCAGAGAA 1260
TTGCTTGAAC CCAGGAGGCG TAGGTTGCAG TGAGTGGAGA TCGTGCCTCT GCAGTCCAGC 1320
CTGGGTGAAA GAGCGAGACT CCGTCTCAAA ATGAAAAAAA AAAAGAAAA CAAGTAGAGA 1380
CTGCAAAAG GGAACAGTAC CGGGAATGTT GGAGAAAAAC ATACTACAAT TAAATCCAAC 1440
ACCCCTGTTG GTCCTGCTAA ATGACAGGCA CTGTGGAAGG TGCTTGGGAC TCAGATAAAT 1500
AAGACAAAGA TCTGCCCATG GAAAGTTCAC GTCTGGACCA TAAGGCATTA GGTTTCATTG 1560
TGAGCTTCCT AGTGGCCAAG GCAAAAAGGA AATAGAATGG TTTAGACAGC TCTCATTGTC 1620
TGATCAAAGG TGTTGAGGCA GAGCACTGAG GAGGGCCTGG AGATAAAGGG TGGGCTGGGG 1680
GTCAGATGCA GTTATCCCTT TGCCGACCCCT TTGTTCCCT TCCTCAGACA CACCTGCCAG 1740
GGTTTGTGGA GCAGGCTGAG GCTCTGAAGG CCAAGGGAGT CCAGGTGGTG GCCTGTCTGA 1800
GTGTTAATGA TGCCTTGTG ACTGGCGAGT GGGGCCGAGC CCACAAGGCG GAAGGCAAGG 1860

TGAGGTGAGG	GGCCTGCAGG	GAGTCAGGAC	CAGGTAGGAT	ATTCTTCTTG	TGACCTCTAC	1920
TTTCTCTGCA	GGTCGGCTC	CTGGCTGATC	CCACTGGGC	CTTTGGGAAG	GTGAGTGTTC	1980
CCCTGACCGC	CACAGGGACA	TGGCGGTGCG	GGGAGCAGTG	GGGGCCCTTG	GCCTCTTCAA	2040
GGATTCTGA	CACTTTCTC	TGTCTCTTCT	TAGGAGACAG	ACTTATTACT	AGATGATTG	2100
CTGGTGTCCA	TCTTGGAAGA	TCGACGTCTC	AAGAGGTAAA	AGTGGAGAGT	CCTCTGTGGA	2160
GAAAGTCCTC	TGTGGGAGAG	AGTCCTCTGT	GGGAGAGAGT	CCTCTGTGGA	GAGGGTCCTC	2220
TGTGGGAAGA	GTCGTCTGTG	GGGGAGATGT	GTGGGAGAGA	GTCCTGTGTG	GGGAGAGTCT	2280
TCTGTAGGGG	AGAGTCCTCT	GGGGAGAGAG	TCCTGTGTGG	GGGAGAGTCC	TCTGTGGGA	2340
GAGTCCTCTG	TGTGGAGAGA	GTCCTGTGTG	GTGGTGAGTC	CTCTGTGGG	GAGAGTCCTC	2400
TGTGGGGGA	GTCCTCTCTG	GAGTTCTCTT	GGGCCCTGG	CTGTTCACTG	CCTGTCTCCA	2460
TGCCCAAGCCT	CCAAGCCCAG	GCTGATGCAG	CTGGCTGGC	CCCTCTTCC	GGCAGGTTCT	2520
CCATGGTGGT	ACAGGATGGC	ATAGTGAAGG	CCCTGAATGT	GGAACCAGAT	GGCACAGGCC	2580
TCACCTGCAG	CCTGGCACCC	AATATCATCT	CACAGCTCTG	AGGCCCTGGG	CCAGATTACT	2640
TCCTCCACCC	CTCCCTATCT	CACCTGCCCA	GCCCTGTGCT	GGGCCCTGC	AATTGGAATG	2700
TTGGCCAGAT						2710

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GCCATCCCAG	CAGTGGAGGT	GTGG	25
------------	------------	------	----

(2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

TTGAACAGCT CTGCCAGGTT CACC

24

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

TGGAGGTGTT TGAAGGGGAG CCAG

24

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

CAGGTTCAAC TTGTTCCCTG GCTC

24

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GGGTATGGGA CTAGCTGGCG

20

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CTGGCCAACA TTCCAATTGCG AG

22

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

ATGTTATGCA ACCCTTGCG ACAC

24

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GTGTTTGAAG GGGAGCCAGG GAAC

24

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

AGAGACAGGG TTTCACCATC TTGG

24